

Segmentasi Karakteristik Pelanggan menggunakan Teknik *Clustering* pada Bisnis *On Demand Service*

Sigit Nugroho¹, Dthomas Hatta Fudholi², Lizda Iswari³

^{1,2,3} Program Studi Teknik Informatika FTI Universitas Islam Indonesia

Jalan Kaliurang Km 14,5 Sleman DI Yogyakarta, telp. +622748984444

e-mail : ¹ 15917121@students.uui.ac.id, ² 085230103@uui.ac.id, ³ lizda@uui.ac.id

Abstrak

Dalam sebuah bisnis, pengetahuan terkait karakteristik pelanggan adalah menjadi mutlak untuk diketahui, untuk mengetahui hal itu dibutuhkan sebuah metode yaitu *marketing mix for big data management framework*. Sumber data dalam penelitian ini diambil dari perusahaan *start up* di Yogyakarta yang memiliki model bisnis *on demand service* berbasis aplikasi, fungsi utama aplikasi ini untuk mempertemukan antara Partner dan Customer terkait kendala aktifitas rutin rumah. Terdapat 5 variabel yang digunakan, yaitu rata – rata rating Partner (skala 1-5), rata- rata selisih jarak order dan workshop, rata – rata selisih waktu order dan pengerjaan, jumlah order, dan rata – rata biaya transaksi. Data diambil dari rentan waktu Mei 2017 sampai dengan Juli 2019, terdapat 1697 jumlah customer, 302 jumlah partner, dan 1483 jumlah transaksi, pengolahan data menggunakan *k-means*, membentuk 4 kluster untuk setiap jenis pelanggan (Partner dan Customer), dimana hasil dari analisis data tersebut memberikan rujukan kepada pihak *marketing* perusahaan dalam pembuatan model pelanggan potensial.

Kata kunci: *marketing mix, data mining, clustering, k-means, segmentasi*

Abstract

In a business, a knowledge related to the characteristics of customers is absolute to be known. In order to know that, a method called *marketing mix for big data management framework* is needed. The source of data in this research are taken from a start-up company located on Yogyakarta. The company's business model is based on demand service, the major function of this application is to connect partners and customers related to the house chore. There are 5 variables which are used, namely: the partner average rating (scale 1-5), the average range of distance the partners take from the workshop, the average gap of time from ordering time and works, the number of orders, and the average of transaction costs. Its data are taken from 2017 may to 2019 july. There are 1,697 number of customers, 302 partners, and 1,483 number of transactions. The data processing is using *k-means* that forms 4 cluster for each customers (partner and customer), which the result of analysis is providing referral to the *marketing firm in making potential customer model*.

Keywords: *marketing mix, data mining, clustering, k-means, segmentasi*

1. Pendahuluan

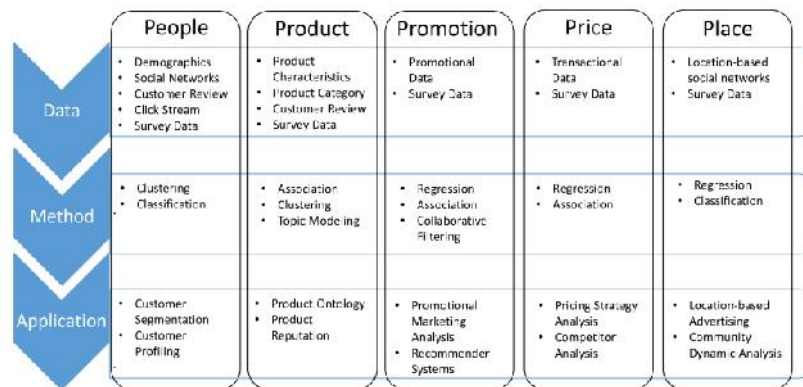
Bisnis *on demand service* (layanan berdasarkan permintaan pelanggan) merupakan bisnis yang sudah tidak asing lagi bagi masyarakat umum belakangan ini, khususnya yang berada di kota – kota besar Indonesia, layanan yang ditawarkan diantaranya seperti, transportasi *online*, pemesanan makanan, pengiriman paket, jasa belanja, jasa kebersihan dan lain sebagainya. GO-JEK Indonesia merupakan sang *unicorn*, menjadi salah satu diantara beberapa yang sejenis lainnya, GO-JEK Indonesia melakukan ekspansi besar - besaran dengan menghadirkan layanan – layanan tersebut secara merata diberbagai lokasi yang dianggap strategis. Melalui semangat yang sama jugalah, tidak sedikit pemain sejenis mencoba ikut serta dalam kegiatan bisnis seperti ini [1]. Sebagai contoh aplikasi A, aplikasi ini merupakan produk *mobile app* dari perusahaan

startup on demand service yang fungsi utamanya adalah mempertemukan antara penyedia jasa (Partner) dan pengguna jasa (Customer), memberikan solusi bagi pelanggannya yang mengalami beberapa kendala dengan aktifitas rutin di rumah mereka, seperti : cuci mobil, bersih – bersih rumah, cuci pakaian, perawatan ac, cuci sepatu & tas [2]. Perusahaan startup ini sudah ada sejak tahun 2016 dan peluncuran pertama aplikasi ini yaitu pada tahun 2017 awal lalu, dari sejak kemunculannya hingga saat ini wilayah operasi masih dalam lingkup Daerah Istimewa Yogyakarta saja, kemudian dalam perjalanannya penyedia jasa yang diajak kerjasama mengalami peningkatan, begitu juga dengan pelanggan yang melakukan order untuk setiap layanan yang ditawarkan.

Dalam sebuah bisnis, strategi pemasaran merupakan kegiatan yang sangat penting untuk dilakukan, karena dengan kegiatan tersebut keputusan – keputusan mengenai sasaran pasar hingga keperluan pemasaran dapat dipetakan secara terperinci [3], seperti halnya yang telah dilakukan oleh tim marketing perusahaan ini, mereka melakukan kegiatan – kegiatan untuk mendukung setiap putusan strategi pemasaran tersebut dengan tujuan bawah pelanggan mereka mendapatkan informasi yang sesuai kaitan layanan yang ditawarkan. Menarik untuk diketahui, sebagaimana yang sudah dijelaskan di atas, bahwa model bisnis (*on demand service*) ini memiliki sudut pandang yang sedikit berbeda dari bisnis lain mengenai posisi mereka dimata pelanggan, karena mereka berada di tengah – tengah di antara ke - 2 (dua) jenis pelanggan yang ada yaitu pelanggan yang menyediakan jasa (customer) dan pelanggan yang menggunakan jasa (partner). Karena posisi itu lah perusahaan ini harus mempertimbangkan bagaimana pendekatan yang mesti dilakukan terhadap ke – 2 jenis pelanggan mereka. Untuk memperkenalkan layanan yang ditawarkan kepada pengguna jasa (customer), mereka melakukan beberapa kegiatan, yaitu diantaranya : kerjasama *endorsement* dengan beberapa *influencer* yang dianggap bisa memberikan pengaruh kepada *follower*-nya, lalu memberikan daftar promo dalam bentuk kode kupon di aplikasi yang dapat dipakai sewaktu customer melakukan order, lalu promo dan informasi terbaru juga disebarkan melalui media sosial mereka (instagram, facebook dan sejenisnya), kemudian mengirimkan pesan masal (sms, line, whatsapp, *email*, *push notification* melalui aplikasi), dan hal tersebut merupakan kegiatan dalam pemanfaatan media promo modern era digital [4]. Sedangkan untuk kegiatan yang dilakukan kaitannya dengan penyedia jasa (partner), tim marketing perusahaan ini memberikan bonus deposit dan setidaknya ada 100 kode kupon promo pada awal kerjasama sampai 1 (satu) bulan pertama (biasanya potongan harga layanan) yang harapannya promo tersebut dapat mereka sebar luaskan kepada pengguna layanan mereka sendiri nantinya.

Dari beberapa putusan strategi pemasaran yang telah mereka ambil, kemudian dijalankan itu sebenarnya harus memiliki alasan yang kuat, tepat dan jelas, sehingga hasilnya dapat memberikan apa yang sebenarnya dibutuhkan oleh pelanggan dan apa yang diinginkan oleh perusahaan [5]. Karena, jika dalam pengambilan keputusannya diputuskan secara sepihak atau bisa dibilang sifatnya hanya menyampaikan informasi searah tanpa melibatkan lebih banyak informasi kebutuhan pelanggan di dalamnya, ditambah lagi dengan minimnya pengetahuan tentang siapa (karakteristik pelanggan baik customer ataupun partner) yang nantinya akan menerima informasi tersebut, maka hal ini bisa jadi memiliki kemungkinan belum atau bahkan tidak menyentuh sasaran yang diinginkan, bisa saja informasi tersebut dianggap tidak relevan kemudian diartikan sebagai spam oleh calon pelanggan [6]. Salah satunya dibuktikan melalui perbedaan data statistik pengguna aktif yang masih mempertahankan aplikasi tetap ter-install - pada smartphone dengan data pengguna (customer & partner) aktif yang ada pada basis data mereka sendiri, bahwa data pengguna aktif yang mereka miliki cenderung lebih banyak dari data pengguna aktif yang ada di google playstore [7]. Sedikit dari bukti data tersebut, sebenarnya sudah dapat memberikan gambaran bahwa ketika sasaran pasar yang diharapkan tidak diketahui dengan tepat (target potensial) pada awalnya, maka akan berdampak dengan terbuangnya waktu dan biaya atas kegiatan yang telah dilakukan itu [8].

Dalam dunia marketing sebenarnya ada pembahasan mengenai obyek – obyek pendukung strategi pemasaran yaitu 4P (*Product, Promotion, Price, Place*) atau yang lebih dikenal dengan *marketing mix* 4P [9], dalam perkembangannya obyek – obyek tersebut kian mengalami perubahan dari yang tadi nya 4P mejadi 5P, 6P, 7P, dan 10P [10], yang dimana menurut hemat penulis, pembahasan dari masing – masing obyek tersebut dapat memberikan gambaran bagaimana cara mengolah beberapa kumpulan data yang ada pada masing – masih obyek untuk kemudian dijadikan sebuah pengetahuan tentang informasi yang belum diketahui sebelumnya, sampai akhirnya bisa memberikan dukungan putusan startegi pemasaran yang didasari dari obyek – obyek data itu sendiri, seperti yang pernah diusulan oleh S. Fan dkk, dalam penelitian meraka kaitan big data analysis dengan judul *Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix*, yaitu *marketing mix framework for big data management* ditunjukkan pada Gambar 1, dalam *framework* tersebut terdapat 5 obyek *marketing mix* (*people, product, promotion, price, place*), dimana masing – masing obyek tersebut menawarkan alur bagaimana dari sekumpulan data, kemudian dengan menggunakan metode tertentu dapat menghasilkan sebuah informasi yang bisa dijadikan landasan putusan strategi pemasaran.



Gambar 1 *Marketing Mix Framework for Big Data Management* [11]

Dari Gambar 1 di atas, terlihat bahwa pemanfaatan teknik – teknik *data mining* menjadi hal yang direkomendasikan untuk digunakan. Karena seperti yang kita ketahui bahwa *data mining* setidaknya memiliki 5 metode yaitu, estimasi, prediksi, kalsifikasi, klastering, dan asosiasi [12]. *Data mining* itu sendiri merupakan sebuah disiplin ilmu terkait dengan metode pengolahan data yang fokus utamanya memberikan sebuah informasi baik itu berupa pola, perilaku dan prediksi yang ada pada data, kemudian dijadikan sebagai pengetahuan yang dapat digunakan untuk mendukung sebuah keputusan [11]. Kemudian mengacu pada apa yang terjadi terkait aktifitas / putusan strategi pemasaran yang telah dijelaskan di atas, yaitu minimnya pengetahuan akan karakteristik pelanggan (*customer & partner*), maka tim *marketing* perusahaan ini memang sudah seharusnya mengetahui dengan detil bagaimana pelanggan mereka selama ini, agar informasi tersebut dapat diketahui jelas diawal, sebelum keputusan strategi pemasaran itu dibuat. Seperti yang pernah diutarakan Wedel & Kamakura pada penelitian berjudul “*Market Segmentation: Conceptual and methodological foundations*”, mereka mengatakan bahwa aktifitas produksi dan pemasaran tidak dapat dilakukan, tanpa mempertimbangkan kebutuhan pelanggan dan menyadari bahwa kebutuhan pelanggan itu memiliki karakter yang berbeda – beda, pendapat ini juga bisa diartikan sebagai salah satu definisi dari segmentasi [13]. Berangkat dari hal tersebut di atas, dan dengan mempertimbangkan salah satu obyek yang direkomendasikan dari *marketing mix framework for big data management*, khususnya untuk mengetahui segmentasi dari pelanggan adalah dengan mengambil obyek *people* sebagai langkah awal untuk memulai pengolahan data yang ada. Dalam obyek *people* tersebut ada 3 (tiga) poin yang harus disiapkan dan atau dipilih, yaitu *data, method, application*. Berikut adalah penjelasannya :

1. **Data**, jika diperhatikan dalam Gambar 1, setidaknya ada 5 jenis data yang harus dipersiapkan, namun dalam penelitian ini hanya akan menggunakan sebagian dari yang direkomendasikan, adapun data yang akan digunakan adalah sebagai berikut :
 - *demographics* : lokasi order customer, lokasi workshop partner,
 - *customer review* : rating customer terhadap partner,
 - *social network*, *click stream* dan *survey* data tidak digunakan, dikarenakan memang data tersebut tidak ada,
2. **Method**, karena data yang dipakai adalah data yang tidak dikelompokkan sebelumnya atau tanpa label, maka metode clustering adalah metode yang sesuai untuk karakter data yang seperti itu [12],
3. **Application**, merujuk dari permasalahan di atas, bahwa segmentasi pelanggan merupakan hal yang harus dibentuk dari proses pengolahan data nantinya.

Dari 3 (tiga) poin di atas, penulis ingin memberikan penjelasan tambahan terkait data lokasi order customer yang digunakan pada bagian demographics. Data lokasi order customer merupakan data yang dihasilkan dari salah satu atribut yang ada pada data transaksi order yang dilakukan oleh customer terhadap partner tertentu untuk kemudian bisa diperoleh data rata - rata selisih jarak antara lokasi order customer dan lokasi workshop partner. Dalam data tersebut juga terdapat beberapa atribut lain, *partner_id* & *customer_id* yang direlasikan melalui data customer & partner untuk kemudian bisa diperoleh data jumlah order yang pernah dilakukan, waktu pemesanan dan pengerjaan layanan order customer untuk kemudian bisa diperoleh data rata – rata selisih waktu pemesanan dan pengerjaan, dan yang terakhir adalah total transaksi dalam bentuk nominal mata uang yang nantinya dapat diperoleh rata – rata total transaksi yang dilakukan.

Jika kombinasi dari atribut – atribut tersebut juga digunakan dalam proses pengolahan data pada penelitian ini, maka setidaknya akan ada kemiripan antara beberapa atribut tersebut dengan kombinasi sumber data yang digunakan oleh N. Morteza dkk, dalam penelitian mereka yang berjudul *An approach to optimised customer segmentation and profiling using RFM, LTV, and demographic features*, dalam penelitian tersebut sumber data yang digunakan berasal dari data transaksi pelanggan dan data profil pelanggan, melalui penelitian tersebut penulis memiliki keyakinan bahwa dengan menambahkan sebagian atribut data yang ada pada data transaksi pelanggan, dapat memberikan informasi serta landasan yang lebih meyakinkan dan mendalam kaitan hasil yang terbentuk dari proses pengolahan data dengan metode kluster nantinya. Lalu melihat dari method dan application yang dipilih di atas, ada sebuah algoritma clustering yang biasa digunakan untuk membentuk segmentasi pelanggan, algoritma tersebut adalah K – Means algorithm. Sebagaimana yang dilakukan oleh Chisty dkk dalam penelitian yang berjudul *RFM ranking – An effective approach to customer segmentation*, para peneliti membandingkan data analisa RFM ke dalam 3 (tiga) algoritma kluster, yaitu K – Means, Fuzzy C – Means, dan RF K - Means, kemudian hasilnya bahwa algoritma K – Means merupakan algoritma yang diusulkan, karena dapat memberikan waktu eksekusi yang lebih sedikit dan juga mengurangi jumlah iterasi [14].

Dari kondisi aktifitas bisnis yang sudah dijelaskan di atas dan dengan berpedoman pada usulan melalui obyek *people* dalam *marketing mix framework for big data management*, bahwa proses *clustering* yang akan dilakukan menggunakan beberapa variabel pembentuk karakter kluster berasal dari data transaksi, data demografi pelanggan (customer & partner), kemudian di aplikasikan ke dalam sebuah algoritma yang disebut K – Means algorithm. Melalui penelitian ini penulis berharap, informasi yang dihasilkan dari proses pengolahan data dan analisis hasil dapat memberikan pengetahuan mengenai karakteristik kelompok order pelanggan baik itu Customer maupun Partner yang akan dicantumkan pada tahap evaluasi pada penelitian ini,

informasi tersebut berupa kelompok order pelanggan dengan riwayat transaksi order seperti apa untuk kemudian dapat dijadikan sebuah model jenis pelanggan yang berpotensi sebagai target pasar oleh pihak marketing perusahaan *on demand service* berbasis aplikasi pada umumnya dan khususnya untuk marketing Perusahaan Aplikasi A ini sendiri.

2. Metode Penelitian

2.1. Pemahaman Bisnis

Pada tahap ini akan dilakukan analisa terkait kondisi bisnis Perusahaan Aplikasi A, mulai dari masalah yang mungkin dihadapi sampai ke penentuan usulan solusi yang ditawarkan, dengan harapan bahwa masalah tersebut dapat diatasi. Sebagaimana yang sudah dijelaskan pada bagian latar belakang penelitian, bahwa kondisi tim marketing Perusahaan Aplikasi A yang terjadi saat ini adalah minimnya pengetahuan terkait informasi karakteristik pelanggan (customer & partner) yang mereka miliki, berangkat dari hal tersebut, melalui penelitian ini kemudian dengan beberapa landasan teori dari kajian terdahulu dan dengan dukungan dari beberapa data yang bisa digunakan, maka penerapan teknik data mining *clustering* untuk model bisnis *on demand service* merupakan tawaran yang diusulkan.

2.2 Kajian Pustaka

Pada tahap ini akan dilakukan studi literatur terhadap penelitian sebelumnya mengenai teori, metode, dan teknik yang relevan untuk mendukung usulan solusi yang ditawarkan pada penelitian ini. Bahwa usulan yang ditawarkan pada kajian terdahulu, memberikan beberapa landasan teori seperti, perlunya segmentasi pelanggan untuk mengetahui karakteristik pelanggan itu sendiri, kemudian penggunaan teknik data mining *clustering* dengan menggunakan algoritma K-Means sebagai tahapan proses pengolahan data.

2.3 Pengumpulan Data

Pada tahap ini akan dilakukan proses pengumpulan data, dimana sumber data tersebut diperoleh dari data sebuah perusahaan startup berbasis aplikasi dengan model bisnis *on demand service* yang berada di kawasan Daerah Istimewa Yogyakarta, dari kumpulan data tersebut perlu dipilih atribut data yang sesuai dengan kebutuhan pada penelitian ini, bahwa data yang digunakan berasal dari data transaksi pelanggan, data demografi pelanggan, dan data layanan, kemudian proses selanjutnya adalah memastikan bahwa semua atribut data tersebut sudah bisa memenuhi kebutuhan variabel / parameter yang akan dipakai pada tahap selanjutnya.

2.4 Transformasi Data

Pada tahap ini akan dilakukan proses transformasi data, dimana data yang sudah terkumpul di awal masih perlu diseleksi sesuai dengan kebutuhan untuk proses pengolahan data nantinya yaitu data yang ada kaitannya dengan demographics, customer review, dan beberapa atribut yang ada pada data transaksi pelanggan, kemudian pada tahap ini juga tidak semua data yang sudah diseleksi dapat langsung digunakan, karena membutuhkan proses tambahan, sebagai contoh untuk mendapatkan data rata – rata selisih jarak antara lokasi order customer dan lokasi workshop partner harus dilakukan proses penghitungan yang dihasilkan dari selisih titik koordinat dari ke dua lokasi tersebut.

2.5 Pengolahan Data

Pada tahap ini akan dilakukan proses pengolahan data dengan menggunakan metode / teknik sesuai dengan usulan solusi yang ditawarkan dalam penelitian ini. Adapun teknik data mining yang dipakai adalah *clustering* dengan menggunakan algoritma K – Means acuan pengolahan data penelitian ini.

2.6 Analisis Hasil

Pada tahap ini akan dilakukan analisa terhadap hasil pengolahan data yang dilakukan, hasil analisa ini diharapkan dapat memberikan informasi berupa pengetahuan mengenai karakteristik pelanggan (customer & partner) berdasarkan beberapa kelompok yang terbentuk.

2.7 Evaluasi

Pada tahap ini akan dilakukan proses evaluasi hasil untuk selanjutnya dituangkan dalam bentuk penarikan kesimpulan, apakah penerapan teknik data *mining clustering* menggunakan K – Means yang digunakan, dan hasil analisa yang didapat dari proses pengolahan data pada penelitian ini, dapat memberikan jawaban terhadap hipotesis awal, yaitu, perlunya sebuah model dari setiap jenis pelanggan baik itu Partner maupun Customer yang memiliki potensi sebagai target pasar oleh pihak marketing perusahaan *on demand service* berbasis aplikasi pada umumnya dan khususnya untuk marketing Perusahaan Aplikasi A ini sendiri.

3. Hasil dan Pembahasan

3.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini terdiri dari data customer, data partner, data transaksi. Semua data yang digunakan dikumpulkan dari rentan waktu Januari 2017 sampai dengan Juli 2019. Data tersebut diambil dari *database* Aplikasi Perusahaan A yang sudah diekstrak dalam format (.xls). Dari hasil data yang sudah dikumpulkan terdapat 1697 jumlah customer, 302 jumlah partner, 1483 jumlah transaksi.

3.2 Transformasi Data

Data yang sudah dikumpulkan kemudian dilakukan penyesuaian terhadap variabel yang digunakan pada proses pengolahan data, sesuai dengan latar belakang penelitian ini bahwa variabel yang digunakan merupakan hasil kombinasi dari beberapa atribut yang ada pada data yang sudah dikumpulkan. Penjelasan mengenai variabel yang digunakan dalam penelitian ini dapat dilihat pada Tabel 1 di bawah ini.

Tabel 1 Variabel Data

No	Kode Variabel	Variabel	Atribut	Sumber Data
1	v1	Rata – rata rating (skala 1-5)	Rating	Partner
2	v2	Rata – rata selisih jarak lokasi order customer dan lokasi workshop partner (Kilo Meter)	Koordinat Pemesanan, Koordinat Workshop Partner	Transaksi & Partner
3	v3	Rata – rata jumlah order yang dilakukan	Customer ID & Partner ID	Transaksi, Customer & Partner
4	v4	Rata – rata selisih waktu order dan pengerjaan (Jam)	Waktu Pemesanan, Waktu Pengerjaan	Transaksi
5	v5	Rata – rata total transaksi order (Rupiah)	Total Transaksi	Transaksi

Untuk menghasilkan data yang siap digunakan dalam pengolahan data sesuai dengan beberapa variabel pada Tabel 1 di atas, maka perlu dilakukan beberapa proses transformasi data dari data mentah menjadi data yang sudah sesuai dengan variabel yang diinginkan. Berikut beberapa proses transformasi data yang dilakukan pada penelitian ini :

1. Untuk mendapatkan data selisih jarak antara lokasi order dan workshop partner, perlu dilakukan *query* terhadap data seperti pada Gambar 5, jenis data lokasi yang ada pada sumber data penelitian ini adalah koordinat *longitude* dan *latitude*. Satuan jarak yang digunakan adalah kilo meter, perlu diketahui bahwa proses *query* ini dilakukan pada tahap ekstraksi data dari database Aplikasi Perusahaan A.
2. Untuk mendapatkan selisih waktu order dan pengerjaan, perlu dilakukan *query* terhadap data seperti pada Gambar 6. Satuan waktu yang digunakan adalah jam, perlu diketahui bahwa proses *query* ini dilakukan pada tahap ekstraksi data dari database Aplikasi Perusahaan A.
3. Bentuk variabel yang digunakan pada penelitian ini merupakan data rata – rata, maka perlu dilakukan pengelompokan (Group By) berdasarkan CustomerID untuk pengolahan data Customer dan PartnerID untuk pengolahan data Partner. Dari hasil pengelompokan yang dilakukan maka akan memiliki pengaruh terhadap jumlah data yang diolah nanti, untuk data Customer dari 1483 jumlah transaksi diperoleh 462 data transaksi order unik berdasarkan CustomerID, sedangkan untuk data Partner diperoleh 81 data transaksi order unik berdasarkan PartnerID (PID). Pada Tabel 2 merupakan data yang sudah dikelompokkan berdasarkan CustomerID (CID).
4. Normalisasi data pada penelitian ini menggunakan rumus MIN MAX *Normalization* yang ada pada MS Excel. Pada Tabel 3 menunjukkan data hasil normalisasi yang dilakukan.

```
SELECT (6371 * acos( cos( radians(a.lokasi_order_customer) ) *
cos( radians(b.lokasi_workshop_partner) ) *
cos( radians(a.lokasi_order_customer) - radians(b.lokasi_workshop_partner) ) +
sin( radians(a.lokasi_order_customer) ) * sin( radians(b.lokasi_workshop_partner) ) ))
as radius_lokasi_pemesanan_dan_workshop FROM transaksi a, partner b
```

Gambar 2 Query selisih jarak order dan workshop

```
SELECT abs(TIMESTAMPDIFF(hour,tanggal_pemesanan, tanggal_pengerjaan))
as selisih_tanggal_pemesanan_dan_jadwal_pengerjaan FROM transaksi
```

Gambar 3 Query selisih waktu order dan pengerjaan

Tabel 2 Data Transaksi Order Group By Partner ID

PID	v1	v2	v3	v4	v5
101	5	1	1,420975	62500	2
159	5	2	3,460051	49167	6
171	5	4	8,212422	200000	1
173	5	8,2	6,386396	54000	5
....

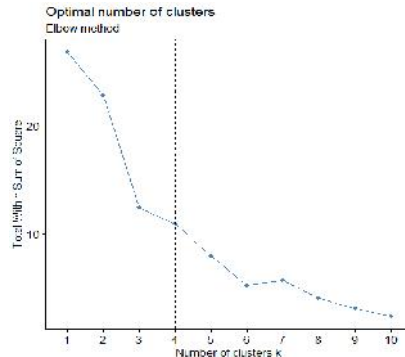
Tabel 3 Hasil Normalisasi Data Partner

PID	v1	v2	v3	v4	v5
101	1	0	0,119316	0,105505	0,010638
159	0	-0,98824	-0,66543	-0,91896	-0,94681
171	2,5	0,047059	0,86693	0,366972	0,010638
173	2,5	0,096471	0,674169	0,099083	0,053191
....

3.3 Pengolahan Data

Pengolahan data ini menggunakan *software* R versi 3.6.1. yang berjalan pada OS Windows. Berikut merupakan langkah – langkah yang harus dilakukan untuk memulai pengolahan data, sebagai contoh langkah – langkah ini merupakan pengolahan *dataset* untuk data Partner saja, sedangkan *dataset* untuk data Customer dilakukan dengan langkah – langkah yang sama seperti pada bagian yang akan dijelaskan berikut ini :

1. *Import* data (.csv) yang sudah dinormalisasikan pada proses transformasi data sebelumnya, dengan perintah “Data=read.csv(file.choose(),header=T)”, kemudian lakukan pencarian lokasi data yang akan diolah,
2. Pastikan *dataset* yang akan diolah sudah ter-*import* ,cara menampilkan *dataset* tersebut gunakan perintah “Data”, karena Data merupakan nama yang sudah di set diawal,
3. Penentuan jumlah K untuk pengolahan data Partner ditunjukkan pada Gambar 7, pada gambar tersebut terlihat rekomendasi K optimal pada titik K=4 maka nilai kluster K yang adalah K=4. Perintah untuk menampilkan data seperti Gambar 7 ditunjukkan pada Gambar 8. Perlu diketahui untuk dapat menjalankan perintah seperti Gambar 5, diperlukan *library* tambahan untuk di – *install* pada R yaitu NBCLust *Library* dengan perintah seperti Gambar 9.
4. *Clustering* K – Means untuk pengolahan data Customer dengan jumlah K=4 yaitu dengan perintah “kmeans(Data,4)”, contoh hasil data *clustering* untk *dataset* data Customer ditunjukkan pada Gambar 10.



Gambar 4 Penentuan jumlah K dengan Elbow

```
# Elbow method
fviz_nbclust(Data, kmeans, method = "wss") +
  geom_vline(xintercept = 4, linetype = 2) +
  labs(subtitle = "Elbow method")
```

Gambar 5 Perintah penentuan jumlah K dengan Elbow

```
pkgs <- c("factoextra", "NbClust")
install.packages(pkgs)

library(factoextra)
library(NbClust)
```

Gambar 6 Perintah *Install* NBCLust *Library*


```

> # Analisis Klasifikasi K-Means
> result_km = kmeans(data, 4)
> result_km
K-means clustering with 4 clusters of sizes 20, 9, 1, 51

Cluster means:
      V1      V2      V3      V4      V5
1 2.25 0.05774118 0.6144793 0.1288172 0.03133596
2 2.50 0.08949042 0.3218031 0.2518849 0.0002742
3 0.00 -0.89833829 -0.6634332 -0.9129296 -0.54680551
4 2.50 0.06605536 0.2512217 0.0418834 0.23660526

Cluster sizes:
[1] 1 1 1 1 2 1 2 2 1 4 1 4 2 4 1 1 4 4 4 4 4 4 4 4 4 4 4 1 1 4 4 4 4 4 2 4 4
[49] 4 4 4 4 4 1 4 1 4 4 4 4 4 4 4 1 4 4 1 1 2 4 1 2 1 4 4 4 4 3 1 4 1 2 1 4

Within cluster sum of squares by cluster:
[1] 4.359032 1.663499 0.000000 4.354333
(between_SS / total_SS) = 0.4145

Available components:
[1] "cluster" "center" "totss" "withinss" "tot.withinss"
[6] "betweenss" "ssb" "ssw" "total" "n"
>

```

Gambar 7 Hasil *Clustering* data Partner dengan K – Means

3.4 Analisis Hasil

1. Data Partner

Berdasarkan proses olah data menggunakan R, bahwa jumlah K yang direkomendasikan menggunakan metode Elbow adalah K=4, berikut merupakan hasil kluster dari data Partner yang sudah dikelompokkan berdasarkan PartnerID di tunjukkan pada Tabel 4.

Tabel 4. Jumlah dan Anggota Kelompok Kluster Partner

Kluster	Jumlah Anggota	Anggota (PID)
1	20	101,173,177,464,849,1208,1631,1786,1770,159,804,854,1160,1275,1325,1427,1539,1752,1871,1912
2	9	348,604,840,1526,622,1330,1477,1747,1879
3	1	171
4	51	525,1201,1421,1527,1625,1913,204,489,600,623,662,664,775,859,861,871,1331,1371,1413,175,337,536,538,603,620,632,644,681,688,726,727,790,822,858,869,875,1025,1034,1065,1110,1135,1190,1205,1348,1410,1571,1577,1621,1724,1868,1977

Dari Tabel 4 di atas dapat dilihat bahwa kluster 4 memiliki jumlah anggota terbanyak yaitu 51 Anggota PID, kluster 1 sebanyak 20 Anggota PID, kluster 2 sebanyak 9 Anggota PID, dan yang terakhir kluster 3 hanya memiliki 1 Anggota PID.

Untuk melihat karakteristik PID dari masing – masing anggota kelompok kluster, maka lakukan pencocokan data PID yang ada pada masing – masing anggota kluster dengan PID yang ada pada data awal sebelum dinormalisasi seperti 2 dengan menggunakan rumus VLOOKUP pada MS Excel, kemudian lihat kembali nilai rata – rata untuk setiap PID pada masing - masing variabel yang membentuk kelompok kluster dengan menggunakan rumus AVERAGE pada MS Excel. Berdasarkan hasil analisis data yang dilakukan di atas, maka pada Tabel 5 menggambarkan bagaimana karakteristik Partner itu terbentuk yang dapat dilihat dari nilai untuk setiap variabel dipisahkan berdasarkan kelompok kluster 1 sampai dengan 4.

Tabel 5 Karakteristik Partner

Kluster	v1	v2	v3	v4	v5
1	5	12	4	48849	10
2	5	4	3	80755	9
3	5	4	8	200000	1
4	5	8	3	39180	24

2. Data Customer

Berdasarkan proses olah data menggunakan R, bahwa jumlah K yang direkomendasikan menggunakan metode Elbow adalah K=4, berikut merupakan hasil kluster dari data Customer yang sudah dikelompokkan berdasarkan CustomerID di tunjukkan pada Tabel 6.

Tabel 6 Jumlah dan Anggota Kelompok Kluster Customer

Kluster	Jumlah Anggota	Anggota (CID)
1	256	1515,1733,1833,1952,637,911,1159,1360,1432,1498,1595,1720,1825,238,396,853, 1954,1907,627,1799,673,1965,574,1948,1562,1902,1391,1543,764,1885,1650,1535,1511,2027,1637,1972,1188,1723,1938,1420,1563,1245,1988,475,1031,1607,1424
2	31	695,552,1457,578,562,788,793,584,1617,248,426,579,745,1936,2057,566,730,1576,554,1411,29,544,608,1277,1466,365,509,592,784,1087,1230,1764,1848,271,457,586,1189,1368,1622,1755,431,
3	143	370,108,618,1118,1508,1478,1790,247,910,211,2011,1272,1464,1086,1375,1763,1455,628,868,275,255,1177,879,728,1795,2010,1271,1384,1549,1537,1037,1185
4	32	

Dari Tabel 6 di atas dapat dilihat bahwa kluster 1 memiliki jumlah anggota terbanyak yaitu 256 Anggota CID, kluster 3 sebanyak 143 Anggota CID, kluster 4 sebanyak 32 Anggota CID, dan yang terakhir kluster 2 sebanyak 31 Anggota CID.

Sama seperti data Partner Untuk melihat karakteristik CID dari masing – masing anggota kelompok kluster, maka lakukan pencocokan data CID yang ada pada masing – masing anggota kluster dengan CID yang ada pada data awal sebelum dinormaliasi dengan menggunakan rumus VLOOKUP pada MS Excel, kemudian lakukan analisis data dengan melihat kembali nilai rata – rata untuk setiap CID pada masing - masing variabel yang membentuk kelompok kluster dengan menggunakan rumus AVERAGE pada MS Excel.

Berdasarkan hasil analisis data yang dilakukan di atas, maka pada Tabel 7 menggambarkan bagaimana karakteristik Customer itu terbentuk yang dapat dilihat dari nilai untuk setiap variabel dipisahkan berdasarkan kelompok kluster 1 sampai dengan 4.

Tabel 7 Karakteristik Customer

Kluster	v1	v2	v3	v4	v5
1	5	5	2	30092	2
2	4	3	2	23776	2
3	5	10	6	55322	2
4	5	29	2	22602	16

3.5 Evaluasi

Dari analisis hasil yang sudah dilakukan, dapat disimpulkan karakteristik dari masing – masing jenis pelanggan dapat dilihat dari Tabel 5 untuk Partner yang terdiri dari 4 kelompok kluster dan Tabel 7 untuk Customer terdiri dari 4 kelompok kluster, dalam tabel tersebut diperlihatkan bahwa setiap kelompok kluster memiliki nilai yang dapat dijadikan sebuah acuan untuk pembentukan sebuah model pelanggan yang dapat dijadikan sasaran target potensial oleh pihak marketing dalam menjalankan aktifitas mereka kedepan.

4. Kesimpulan

Berdasarkan pembahasan hasil yang telah dilakukan pada penelitian ini, maka dapat disimpulkan bahwa :

1. Berdasarkan pengolahan data dan analisis hasil kluster menggunakan K – Means dan optimasi jumlah K dengan metode Elbow, dihasilkan 4 kelompok kluster untuk masing – masing jenis pelanggan yaitu Partner dan Customer, dimana setiap kelompok memiliki karakteristik pada setiap variabel yang ada,
2. Melalui hasil yang diperoleh dari semua tahapan yang sudah dilalui pada penelitian ini, maka hasil tersebut dapat dimanfaatkan oleh pihak marketing jenis usaha *On Demand Service* khususnya untuk membuat sebuah model pelanggan yang potensial kedepan.

Daftar Pustaka

- [1] R. Eka, “Penerimaan Masyarakat Indonesia terhadap Layanan On-Demand | Dailysocial,” 2017. [Online]. Available: <https://dailysocial.id/post/penerimaan-masyarakat-indonesia-terhadap-layanan-on-demand>. [Accessed: 13-Mar-2019].
- [2] Wellhos, “Official Website Perusahaan Aplikasi A,” 2017. .
- [3] W. A. Triyanto, “ALGORITMA K-MEDOIDS UNTUK PENENTUAN STRATEGI PEMASARAN PRODUK,” *J. SIMETRIS*, vol. 6, 2015.
- [4] A. Yulistara, “5 Strategi Jitu untuk Promosi Bisnis di Era Digital,” *CNBC Indonesia*, 2018. [Online]. Available: <https://www.cnbcindonesia.com/entrepreneur/20180402121905-25-9338/5-strategi-jitu-untuk-promosi-bisnis-di-era-digital>. [Accessed: 13-Mar-2019].
- [5] C. Lin and C. Hong, “Development of a marketing information system for supporting sales in a Tea-beverage market,” *Expert Syst. Appl.*, vol. 36, no. 3 PART 1, pp. 5393–5401, 2009.
- [6] P. Kotler, H. Kartajaya, and I. Setiawan, *Marketing 4.0 : Bergerak dari Tradisional ke Digital*. 2019.
- [7] Google Delveloper, “Dashboard Google Developer Perusahaan Aplikasi A,” 2019. .
- [8] A. Ikhwan, D. Nofriansyah, and Sriani, “Penerapan Data Mining dengan Algoritma Fp-Growth untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus Kampus STMIK Triguna Dharma),” *Saintikom*, vol. 14, no. 3, pp. 211–226, 2015.
- [9] P. Kotler and G. Armstrong, *Principles of Marketing 14 | E*. 2012.
- [10] H. Misbah and L. F. Latif, “‘ Marketing Mix ’: Tinjauan Semula Marketing Mix : a Revisit From Islamic,” *J. Muamalat Islam. Financ.*, vol. 14, no. 1, pp. 87–105, 2017.
- [11] S. Fan, R. Y. K. Lau, and J. L. Zhao, “Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix,” *Big Data Res.*, vol. 2, no. 1, pp. 28–32, 2015.
- [12] R. S. Wahono, *Kuliah 10 Menit tentang Data Mining*. 2014.
- [13] W. Kottier, “Marketing strategy for online SMEs,” 2017. [Online]. Available: https://essay.utwente.nl/71976/1/Kottier_MA_BMS.pdf.
- [14] A. J. Christy, A. Umamakeswari, L. Priyatharsini, and A. Neyaa, “RFM ranking â€“ An effective approach to customer segmentation,” *J. King Saud Univ. - Comput. Inf. Sci.*, 2018.